

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
27 December 2001 (27.12.2001)

PCT

(10) International Publication Number  
**WO 01/99109 A1**

(51) International Patent Classification<sup>7</sup>: **G11B 20/10**

(21) International Application Number: PCT/KR01/00975

(22) International Filing Date: 8 June 2001 (08.06.2001)

(25) Filing Language: Korean

(26) Publication Language: English

(30) Priority Data:  
2000/31255 8 June 2000 (08.06.2000) KR  
2000/31931 10 June 2000 (10.06.2000) KR

(71) Applicant (for all designated States except US):  
**MARKANY INC.** [KR/KR]; Ssanglim Bldg. 10Fl.,  
151-11, Ssanglim-dong, Chung-gu, Seoul 100-400 (KR).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **SHIN, Seung-Won**  
[KR/KR]; 610-207, Shinnae APT., Shinnae 1-dong,  
Jungang-gu, Seoul 131-131 (KR). **KIM, Jong-Woen**  
[KR/KR]; 201 ho, 112-39, Seongnae 2-dong, Seong-  
dong-gu, Seoul 134-032 (KR). **LEE, Han-Ho** [KR/KR];

101 ho, Ga-dong, Shinjin, Myunmok 4-dong, Jun-  
grang-gu, Seoul 131-204 (KR). **PARK, Chang-Mok**  
[KR/KR]; 220-20, Shiheung 5-dong, Gumcheon-gu, Seoul  
153-035 (KR).

(74) Agent: **KOREANA PATENT FIRM**; Dong-Kyong Bldg.  
824-19, Yoksam-dong, Kangnam-ku, Seoul 135-080 (KR).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,  
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,  
HU, ID, IL, IN, IS, JP, KE, KG, KP, KZ, LC, LK, LR, LS,  
LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO,  
NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR,  
TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

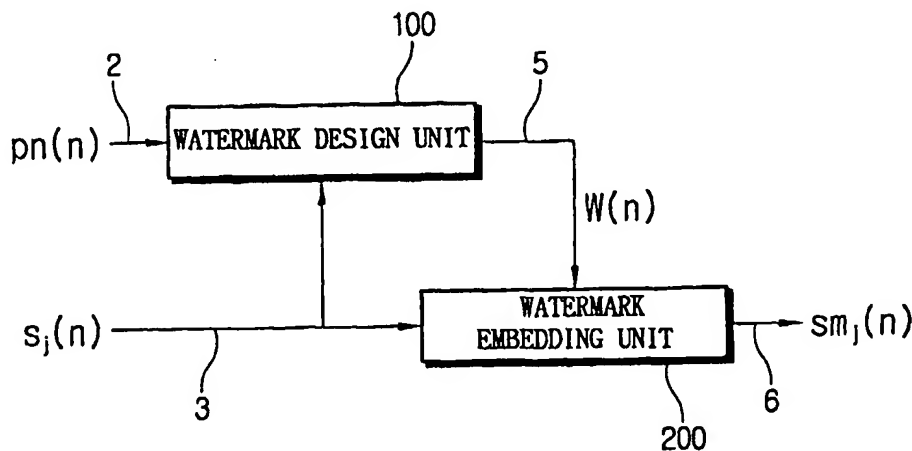
(84) Designated States (*regional*): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian  
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European  
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,  
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

[Continued on next page]

(54) Title: WATERMARK EMBEDDING AND EXTRACTING METHOD FOR PROTECTING DIGITAL AUDIO CONTENTS  
COPYRIGHT AND PREVENTING DUPLICATION AND APPARATUS USING THEREOF



(57) Abstract: In embedding a watermark in a digital audio contents, a model using a minimum audible limit of the digital audio contents is generated and a series of pseudo random number is masked to generate the watermark to be inserted in the digital audio contents. In order to extract the watermark embedded in the digital audio contents, an auditory psychological model having a minimum audible limit, which depends on the characteristic of frequency generated from the digital audio contents, is formed to generate the watermark from a series of pseudo random number. After generating the watermark, the length of the watermark is adjusted according to the relationship of a signal processing between the adjacent digital audio contents. By measuring the correlation between the watermark and digital audio contents, it is possible to detect whether the watermark is inserted or not.



— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

**WATERMARK EMBEDDING AND EXTRACTING METHOD FOR  
PROTECTING DIGITAL AUDIO CONTENTS COPYRIGHT AND  
PREVENTING DUPLICATION AND APPARATUS USING THEREOF**

**BACKGROUND OF THE INVENTION**

**Field of the Invention**

The present invention relates to watermark embedding and extracting method in/from digital contents and, in particular, a method and device for embedding  
10 watermark in digital audio contents in which watermark is able to be embedded even in  
a case the digital audio contents have its periodic characteristics and are susceptible  
only to a minor change in quality, and a method and device for extracting watermark  
from the watermark-embedded digital audio contents even in case the original contents  
are in the state of distortion from the attacks of the watermark-embedded digital audio  
15 contents by signal processing.

**Description of the Related Art**

Embedding/extracting watermark in/from the digital audio contents is  
suggested in "Digital Watermarks for Audio Signals" (Laurence Boney, A.H. Tewfik,  
20 and K.N. Hamdy, in Proc, 1996 IEEE Int. Conf. Multimedia Computing and Systems,  
and Hiroshima, Japan, June 17-23, 1996, pp. 473-480). Embedding watermark using  
psycho-acoustic model is suggested by Swanson et al. in "Robust Data Hiding for  
Images" (M.D. Swanson, B. Zu, and A.H. Tewfik, in Information Hiding: Second Int.  
Workshop (Lecture Notes in Computer Science), vol. 1525, D. Aucsmith, Ed. Berlin,

Germany: Springer-Verlag, 1998, pp. 169-190).

The methods for embedding watermark disclosed in Laurence et al. and Swanson et al. are generally known technology. The technology disclosed in Laurence et al. should analyze audio contents per frame and thus a large amount of computation and memory are necessary. Thereby, it is difficult to be practically commercialized.

Further, a method for detecting watermark using correlation widely used in a research relating to the watermarking is the same as a method used in a research of "A Secure Robust Watermark for Multimedia," (I. J. Cox, J. Kilian, T. Leighton, and T. Shamoon in R.J. Anderson, Ed., "Information hiding: First International Workshop," in Lecture Notes in Computer Science, vol. 1174, Berlin, Germany: Springer-Verlag, 1996, pp. 183-206) of Cox et al. and "Secure Spread Spectrum Watermarking for Images, Audio, and Video" (I. J. Cox, J. Kilian, T. Leighton, and T. Shammon in Proc. IEEE Int. Conf, Image Processing (ICIP'96), Lausanne, Switzerlandm Sept. 16-19, 1996, pp. 243-246). However, such a method will cause a problem in which the ratio of detecting watermark just with a simple correlation after the contents are attacked is significantly lowered due to a periodic characteristic of the audio contents.

Further, the digital audio contents have a very close relation with a magnitude of the surrounding signal in respect of its characteristics. Hence, the form of wave of an audio signal itself is comprised of the sum of sine waves (sine curve and cosine curve) having different frequencies to each other. That is, since the sine waves are periodic functions and have periodic characteristics so that a watermark is embedded with such a strength as not to damage quality of the digital audio signals, a frequent concealment of the correlation information occurs under the periodic characteristics of the audio and thus is not perceived.

As an alternative of resolving the above-mentioned problem, International Patent Laid-Open No. WO9803014 suggests an improvement to strengthen the information of watermark by filtering the watermark-embedded digital contents and watermark signal with a predicting filter before finding the correlation function between the digital contents signal and watermark. The above-mentioned International Patent Laid-Open discloses a method for detecting the watermark which includes a step of evaluating the correlation produced by correlating the digital contents and watermark embedded in the information signal. In this case, the watermark-embedded digital contents and watermark are filtered with the predicting filter according to the characteristics of the digital contents, and the resultant correlation is applied to the filtered signal and filtered watermark.

However, the above International Patent Laid-Open explains an embodiment for the object of an image, which is less periodical compared to that of a filtering which requires a considerable time for computation in filtering the watermark-embedded digital contents by a predicting filter, and thus it is impossible to apply to a system requiring a real-time application.

### SUMMARY OF THE INVENTION

An object of the present invention in order to resolve problems as mentioned above is to provide a method for real-time embedding and extracting a watermark without affecting the quality of the audio and a device using the same.

Another object of the present invention is to provide a method for embedding/extracting a digital watermark and a device using the same, which is able to make the whole process of real time embedding/extracting watermark and can be

applied to a portable device such as a MP3 player.

Another object of the present invention is to provide a method for embedding a digital watermark and a device using the same which designs and embeds a watermark using a domain undistinguishable by human audibility.

5       The embedding watermark as mentioned above is done in a domain of a time-spatial space, and a masking method using a digital psycho-acoustic model suggested in the present invention, i.e. audio absolute threshold curves of hearing before the step of embedding watermark is provided.

Another object of the present invention is to provide a method that is able to  
10       detect an embedded watermark excluding the periodic characteristics essential to the digital audio and a device using the same.

In order to accomplish the above objects, the method for embedding watermark in the digital audio contents including the step of generating the digital watermark by filtering a predetermined length of pseudo random sequence with an audio absolute  
15       threshold of hearing in order not to affect the audio characteristics of the digital audio contents; and the step of embedding the digital watermark in the digital audio contents.

The method for extracting watermark from the digital audio contents includes a step of generating watermark from pseudo random number sequence by forming a psycho-acoustic model having an audio absolute threshold of hearing  
20       according to the characteristics of the frequency generated from said digital audio contents; a step of adjusting the length of said watermark by detecting whether or not the signal is processed with regard to the corresponding digital audio contents from the digital audio contents which are adjacent to each other after said step of generating watermark; and a step of detecting the embedding of watermark by measuring the

correlation between said watermark and said digital audio contents.

The above step of detecting the watermark in the method for detecting a watermark according to the present invention, first, decides on how much a signal processing performed on the digital contents affects the strength and delay of watermark.

5 In particular, it is very critical to make sure that a counter-measure be made when a portion of the watermark disappears, which enables a successful detection of watermark.

Secondly, the signal of watermark embedded in the digital audio data must be strengthened. If watermark affects a tone quality or tone color of the digital contents,  
10 it is meaningless even though the watermark can be successfully discriminated. In conclusion, the weak watermark signal must be strengthened, since the watermark signal can not be but be weakly embedded into a signal.

Thirdly, only the identifying signal of watermark must be selectively extracted through the correlation function between watermark-embedded digital signals and  
15 watermark. In a process of extracting the watermark signal, the correlation between the signal contents and watermark signal inevitably makes an accurate extraction difficult. Hence, a method of extracting watermark using a signal processing method which removes the periodic characteristics of an audio from the signal generated in the correlation function between the embedded watermark and watermark signal.

20 The present invention provides a method for extracting watermark embedded in the signal with a strong periodicity such as the digital audio in a high-speed manner and with effectiveness. It is possible to extract the embedded information by this extraction method as long as deterioration in the quality of the audio contents maintains the commercial value of audio even after the attack of an audio signal processing

(analog to digital transformation, transformation in a sampling ratio, transformation in a linear speed, lossy compression , echo hiding).

More detailed explanation on watermark embedding and extracting method for protecting digital audio contents copyright and preventing duplication and apparatus  
5 using thereof is presented hereinbelow in reference to the attached drawings.

### BREIF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a schematic block diagram showing constitution of a device for embedding the digital watermark according to the present invention.

10 Fig. 2 is a schematic block diagram showing constitution of the watermark design unit in Fig. 1.

Fig. 3 is a schematic block diagram showing constitution of the watermark embedding unit in Fig. 1.

Fig. 4 is a schematic block diagram showing constitution of a device for  
15 extracting the digital watermark according to the present invention.

Fig. 5 is a schematic block diagram showing constitution of the watermark design unit in Fig. 4.

Fig. 6 is a detailed block diagram showing the signal variation sensing unit in Fig. 4.

20 Fig. 7 is a detailed block diagram showing the watermark detecting unit in Fig. 4.

Fig. 8 is a schematic block diagram showing constitution of the watermark information authentication unit which extracts the information and whether or not the watermark is embedded from the watermark detecting unit in Fig. 4.



Fig. 9 is a graph showing audio absolute threshold curves of hearing.

Fig. 10A is a graph showing the result obtained by using the conventional method for calculating the information of correlation, and Fig. 10B is a graph showing the result obtained by using a method for detecting a watermark suggested in the present invention.

Fig. 11 is a graph showing the result of detecting watermark after lossy compression.

Fig. 12 is a graph showing the result of extracting watermark in case of 4-bit information.

#### DETAILED DESCRIPTION OF PREFERRED EMBODIMENT

First, a method for embedding the digital watermark and a device using the same relating to the digital audio contents will be explained referring to the drawings from a method for embedding/extracting a watermark and device using the same relating to the digital audio contents according to the present invention.

Fig. 1 is a schematic block diagram showing constitution of a device for embedding the digital watermark according to the present invention. In Fig. 1, numeral 100 indicates a watermark design unit for designing a digital watermark in order to minimize influence on a quality of audio considering the characteristics of the digital contents, and numeral 200 indicates a watermark embedding unit for embedding the digital watermark generated from the watermark design unit into the digital contents.

The watermark design unit 200 designs watermark  $w(n)$  by using input pseudo-random sequence  $pn(n)$ , so as not to minimize influence on an audibility of the corresponding digital contents  $s_j(n)$ . The designed watermark is adjusted suitably to

the amount of energy of the digital contents  $s_j(n)$  input by the watermark embedding unit 200 and is added to the digital contents, thereby the watermark-embedded digital contents  $sm_j(n)$  are provided.

Fig. 2 is a schematic block diagram showing constitution of the watermark design unit in Fig. 1 and illustrates the constitution for designing the watermark which can be harmonized with the digital contents  $s_j(n)$  from the input pseudo-random sequence  $pn(n)$  2. The watermark design unit 100 includes a psycho-acoustic (visual) model 110, a perceiving limit bandwidth filter (audio absolute threshold filter of hearing) 120, and a main signal copy filter 130 for copying the characteristics of the digital contents as the object of embodiment..

The pseudo-random sequence  $pn(k)$  which is put into the watermark design unit 100 is transformed into the filtered signal  $x(n)$  by using a model provided by the psycho-acoustic (visual) model 110 through the perceiving limit bandwidth filter 120. A signal  $wn(n)$  is provided through a main signal copy filter 130 by using a composed filtering coefficient ( $a_i$ ) which represents general characteristics of the digital contents so that the filtered signal  $x(n)$  makes a watermark with a property similar to the characteristics of the digital contents. At this time, the perceiving limit bandwidth filter 120 is a filter with audio absolute threshold curve to which human listens the music referring to a curve computed by a statistical analysis in regard to the audio signal.

Usually, the audio absolute threshold of hearing is the minimum level of audio perceivable by hearing when it is silent, and means the limit of noise perceivable by hearing when it is silent. Graph A among the graphs shown in Fig. 9 shows the audio absolute threshold of hearing when it is silent, which varies according to the frequency of sound (high and low of sound). That is, even the same volume of audio can be

heard or can not according to the frequency.

However, what is intended by the present invention is directed not to the audio absolute threshold of hearing as a general concept described in the above but to the audio absolute threshold of hearing shown usually when the music is played. That is, like graph B of Fig. 9, it means the limit of noise perceivable by hearing at least when the music is played as a result from the analysis of the frequency distribution regarding the different genre of music.

On the supposition of such basic term, the above process presented in an equation is as follows:

$$x(n) = \sum_k apm(k) pn(n-k) \quad (1)$$

$$w(n) = \sum_i a_i x(n-i) \quad (2)$$

If (2) is substituted with (1),

$$w(n) = \sum_i a_i \sum_k apm(k) pn(n-k) \quad (3)$$

Wherein,  $a_i$  can be obtained by extracting the characteristic of the digital contents using an auto-regression (AR) model, moving-average (MR) model, autoregressive moving-average (ARMA) model, and etc.

Further,  $apm(k)$  is a coefficient of the audio absolute threshold filter of hearing of the digital audio used as the psycho-acoustic model 110 and is revised by analyzing the frequency characteristic generated from an audio signal to use publicly the audio absolute threshold of hearing of the psycho-acoustic model in the digital audio. Graph B shown in Fig. 9 is made by analyzing a total of 20 tunes of various types of the audio frequency and shows the audio absolute threshold curve. The graph is presented as the following formula,  $Y=P(1)x^7+P(2)x^6+...+P(7)x+P(8)$ , wherein P means the value of

coefficient showing the limited volume according to the frequency band and has the following values: [1.056801742606838e-026, -7.214332602361358e-022, 1.809126572761631e-017, -1.941502598267307e-013, 5.982813623951169e-010, 4.211560372433627e-006, -3.420594737587419e-002, 8.533065083348841e+001].

- 5 The above coefficient values are given for exemplification and depend on changes according to the characteristics of the audio.

When the pseudo-random number sequence  $pn(k)$  is embedded into the digital contents, the psycho-acoustic model 110 changes the quality of the contents audibly to prevent from being discriminated against the signal wherein a watermark is not  
10 embedded. Generally, human's ears that perform the same role as a spectrum analyzer (a device analyzing frequency) react very sensitively to the audio generated in the frequency band of 3kHz~5kHz, and are scarcely able to discriminate the audio of the frequency band of more than 10kHz.

The psycho-acoustic model 110, which is made by using the frequency  
15 analyzing characteristics of such ear masks the frequency of the pseudo-random number sequence with the audio absolute threshold curve of hearing in order to minimize deterioration in the quality of the audio by the embedded information when embedding the pseudo-random number sequence into the audio.

The masking controls to weaken a signal of a sign sensitive to human's ears  
20 and embeds an original or a larger size of component in an insensitive frequency bandwidth.

Further, the psycho-acoustic model 110 makes the digital contents and embedded watermark signal to maintain the quality of the contents, and intensifies the characteristics of them under signal. Such characteristics can help detect the watermark

by a component of the other parts, even if a part of the frequency component is removed since the size of signal is adjusted according to the characteristics of the audio frequency when masking the pseudo-random number sequence by using the psycho-acoustic model.

5           Moreover, since the psycho-acoustic model 110 is a device used in making a file transformation through most of the lossy compression (MP3, AAC, WMA), it minimizes loss of the watermark information. Omission of the psycho-acoustic model 110 is possible when a very simple processing of watermark is necessary. Omission of this process is possible when there is a limit in the amount of computation although this  
10   process is to change the characteristics of watermark and audio signal into the same type as possible.

The watermark signal  $w(n)$  5 generated from the watermark design unit 100 is embedded into the watermark embedding unit 200, and then the signal wherein the watermark is embedded into the digital contents is finally output. Such will be  
15   explained referring to the constitution of watermark embedding unit 200 shown in Fig. 3. The watermark embedding unit 200 comprises a gain calculator 210 for calculating the strength  $g$  (9) of the watermark which is embedded by measuring the amount of energy of the signal  $s_j(n)$  (3) of the digital contents, a watermark intensity adjustor 220 for adjusting the strength of watermark  $w(n)$  5 which will be embedded according to  
20   the strength obtained by the gain calculator 210, and a watermark signal adder 230 for outputting the digital audio contents  $sm_j(n)$  6 where a watermark is embedded by adding and combining the watermark  $g.w(n)$  10 to the signal of the digital contents.

The gain calculator 210 measures the amount of energy of the digital contents by using volume or the characteristics of the frequency distribution. The volume of

audio is that of the music sound wherein the maximum value per frame is a reference and the frequency distribution is in a range of the frequency expressed by an audio and measures the amount of energy according to whether the frequency of audio is distributed equally or partially in the whole width band.

5 More specifically, the matter considered when deciding the strength of embedding watermark is the volume of audio, and the damage of the sound quality can be reduced only if the strength of watermark is changed in proportion to the volume of audio. For this, the audio data with the frame size is brought and then the frame size is divided by N. The reason for its division by N is to prevent the strength of watermark  
10 from growing much bigger since quite a big size of the frame results in a much difference in the sound volume of audio. For example, the frame size of 1470 is divided by the sub frame of 147 and then the strength of watermark is decided on the basis of the maximum value of each sub frame size.

Further, since audio differs in the distribution range of the frequency by genre,  
15 if the strength of watermark is decided using only the size of the sound volume, it [said watermark] is embedded in some music too minutely or too strongly. Hence, the strength of watermark should be decided considering genre of the music and using the distribution range of the frequency.

In the watermark embedding unit 200 having such constitution, the process  
20 presented in a formula is as follows:

$$sm_j(n) = s_j(n) + g \cdot w(n) \quad (4)$$

The following formula can be obtained from formulas 1 to 3:

$$s_j = \{s(jN + n), n = 0, 1, \Lambda, N - 1, j = 0, 1, \Lambda, J - 1\} \quad (5)$$

$$sm_j(n) = s_j(n) + g \cdot \sum_i a_i \sum_k apm(k) pn(n-k) \quad (6)$$

$$sm_j = \{sm_1, sm_2, \dots, sm_J\} \quad (7)$$

If watermark is embedded in the entire range of the digital contents, the digital contents  $sm_j$  wherein watermark is embedded is finally obtained. When the digital contents are divided by the size of  $N$  and is composed of the number  $J$ ,  $N$  is the length of the pseudo-random number sequence, and  $J$  is an integer obtained by dividing the length of the digital contents is divided by  $N$ . In the present application, the audio signal is divided into a frame with the size  $N$  and the watermark is embedded in each frame. Since watermark uses the pseudo-random number sequence, the size of watermark is the frame size of audio of  $N$  and thus  $J$  is the number of the frame.

A method of extracting watermark and device using the same according to the present invention will be explained with reference to the drawings.

Fig. 4 is a schematic block diagram showing constitution of a device for extracting the digital watermark according to the present invention. The device for extracting watermark in Fig. 4 includes a watermark design unit 300 similar to that used in the prior watermark embedding device, a signal variation sensing unit 400 for detecting whether there is a change in the digital contents, a watermark detecting unit 500 extracting watermark by comparing the watermark in the watermark design unit 300 with the watermark embedded in the digital contents, and a watermark information authentication unit 600 for performing the authentication operation of the corresponding information according to the result from extraction.

In order to extract watermark from the input digital contents, the watermark design unit 300 generates watermark  $w(n)$  by using the generated pseudo-random

number sequence and transmits it to the signal variation sensing unit 400. In case of the digital contents, watermark  $w(n)$  5 is transmitted in the same form or in an order changed form during a transmitting process compared to the time when watermark is embedded. That is, in case that the signal processed in the digital contents removes a part of the watermark information, there should not be any difficulty in extracting (detecting) watermark. For this, a process of re-sampling watermark is executed.

For such process, the signal variation sensing unit 400 decides on whether the signal processed in the digital by using the contents signals  $sm_j(n)$  and  $sm_{j+1}(n)$  which are adjacent contents to each other affects the watermark detecting unit 500. In case that the signal processing affects the watermark detecting unit 500, for example, the proceeding speed of the digital contents is arbitrarily changed or a part of signal is removed, watermark  $w(n)$  (5) is re-sampled, and then the sampled watermark  $w(n)$  (6) is delivered to the watermark detecting unit 500. If there is no change in the digital contents, there is no change in the length of watermark. However, if a change happens,  $n$ , the length of watermark is re-sampled to  $\bar{n}$ . On the supposition that no change happens in the digital contents, the explanation will be made with the following formulas.

The watermark detecting unit 500 calculates amount  $c(n)$  of the correlation information between watermark  $wr(n)$  designed in the watermark design unit 300 and the watermark  $wr(n)$  in the digital contents signal  $sm_j(n)$  wherein watermark is embedded, and from the calculated correlation information, said unit extracts whether a watermark is embedded and extracts the embedded information through the watermark information authentication unit 600. The above-mentioned explanation separately represents the watermark detecting unit 500 and the watermark information



authentication unit 600, but only the watermark detecting unit 500 can be constituted. The process of extracting from the digital contents where such watermark is embedded will be more specifically explained referring to Fig. 5 to Fig. 7.

Fig. 5 shows the watermark design unit for extracting watermark similar to that shown in Fig. 2. Only, it is not necessary to adjust the intensity of watermark embedded in the digital contents in the watermark design unit 300 which is included in the extracting device. Thus, the watermark design unit comprises the psycho-acoustic model 310 and the perceiving limit bandwidth filter 320. The motion of the above watermark design unit 300 is similar to the constituent shown in Fig. 2. Therefore, the detailed explanation will be omitted.

Fig. 6 is a detailed block diagram showing the constitution of signal variation sensing unit 400 in Fig. 4. The signal variation sensing unit 400 in Fig. 6 comprises a watermark restoring filter 410 for its inverse transformation since signal variation sensing unit 400 is transformed into the characteristics of the digital contents when watermark is embedded into the digital contents, a correlation information calculating unit 420 for measuring the auto correlation information of the digital contents where watermark is embedded, an ensemble average calculating unit 430 to strengthen the correlation information using an ensemble average considering that watermark is equally included in the digital contents, and a resampling unit 440.

Signal variation sensing unit 400 receives two digital contents signals,  $sm_j(3)$  and  $sm_{j+1}(4)$  which are adjacent to each other as input signals to detect a change in the digital contents. When the above two signals are input, the watermark restoring filter 410 executes a process of transforming inversely itself again since it is transformed into the characteristics of the digital contents when watermark is embedded into the digital

contents, i.e., a process of restoring to the status prior to embedding a signal of watermark embedded in the digital contents. However, although watermark is not designed elaborately or restored during the extracting process with elaboration, it does not affect the extracted result much, thus, this process can be omitted.

5        When the restoring process is completed, the auto correlation information of the watermark embedded digital contents is determined through the measurement without using watermark to decide on how the signal of the digital contents changes. The correlation information calculating unit 420 extracts the correlation information between the digital contents signals of  $sm_j(n)$  and  $sm_{j+1}(n)$ . After calculation, since  
10        watermark is equally included in the digital contents, the ensemble average calculating unit 430 strengthens the correlation information using the ensemble average.

Finally, considering the correlation between watermark  $w(n)$  input from the watermark design unit 300 and watermark embedded in the digital contents, if no change in the digital contents is found, there is no change in the length of watermark in  
15        case of no change in the digital contents. However, when there is a change, the resampling unit 440 outputs watermark  $wr(n)$  of which the length is resampled from  $n$  to  $\tilde{n}$ .

A process for removing unnecessary information generated through the correlation of watermarks made by the above process will be expressed using equations  
20        as follows:

Watermark-embedded digital signals  $sm_j(n)$  3 and  $sm_{j+1}(n)$  4 adjacent to  $sm_j(n)$  3 output from watermark restoring filter 410, i.e. inverse auto regression filter, are presented in the form of  $x(n)$  9 and  $y(n)$  10 as follows:

$$x(n) = \sum_i b_i sm_j(n-i) \quad (8)$$

$$y(n) = \sum_i b_i sm_{j+1}(n-i) \quad (9)$$

Here, since  $sm_j(n) = w(n) + s_j(n)$ , the following formula is obtained if this is put into Equation (8):

$$x(n) = \sum_i b_i (w(n-i) + s_j(n-i)) = \sum_i b_i w(n-i) + \sum_i b_i s_j(n-i) \quad (10)$$

Here, if  $s_j(n)$  is random and  $s_j \perp s_i$ , if  $i \neq j$ , it is expressed as the following watermark information and errors:

$$x(n) = w(n) + e_j(n) \quad (11)$$

$$y(n) = w(n) + e_{j+1}(n) \quad (12)$$

The two signals  $x(n)$  and  $y(n)$  obtained from the above are input into the correlation information calculating unit 420, and then the amount  $s_3(n)$  of the correlation information is calculated. After calculation of the correlation information, the ensemble average is determined by using the correlation information in the ensemble average calculating unit 430. The ensemble average divides the audio signal by the size of  $N$ , which is called a frame each by each, and said each frame is added and averaged to an average frame, which is called an ensemble average, which is expressed as the following formula:

$$\begin{aligned} c(n) &= \sum_k x(n) y(n+k) = \sum_k (w_D(n) + e_j(n)) (w_{D+1}(n+k) + e_{j+1}(n+k)) \\ &= \sum_i [w(n)w(n+i) + \varepsilon(n)] \quad (e_j \perp e_i = 0) \end{aligned} \quad (13)$$

If the signal processing on the digital contents affects only the strength, the signal  $c(n)$  calculated herefrom can obtain a peak of the maximum value (or minimum value) per length of the watermark used in its early embedding. For example, if there

is no change in the length of the digital contents such as change due to an attack of weakening the watermark information through noise or filtering, i.e., a simple attack of noise or signal processing, a peak appears in a period of the length (N) of watermark used when the watermark signal is embedded.

5           However, if a signal processing increases or decreases the audio length, a peak appears per length of  $\tilde{n}$ . Therefore, the watermark signal  $w_r(n)$  is obtained by resampling the watermark  $w(n)$  using such information. It is possible to use an known method for resampling and; when a high-speed processing is necessary, it is better to use a spline extrapolation.

10           For example, there could be an attack (representatively, "pitch shift") which gives a change in the length of the digital contents preventing watermark from being found by embedding different sample signals amid the digital contents. In this case, the length of frame is changed. For example, if the length of the digital contents increases by 10% and the size of frame used in embedding watermark is  $n$ , the size of  
15   frame after being attacked is  $N*1.1$ , and thus the peak value of the correlation which confirms the existence and non-existence of watermark during the process of extracting watermark appears in every  $n*1.1$ .

Fig. 7 is a block diagram showing the detailed constitution of the watermark detecting unit 500 in Fig. 4 and shows the constitution of extracting for determining the  
20   amount of the correlation information which is consistent with watermark in the correlation information between the watermark-embedded digital contents and watermark. The watermark detecting unit 500 in Fig. 7 includes a watermark restoring filter 410 for executing the same function as that of the watermark restoring filter 410 used in the signal variation sensing unit 400 in Fig. 6, an ensemble average calculating

unit 520 for strengthening the intensity of the watermark signal in the restored digital contents, a correlation information calculating unit 530 for calculating the correlation information between the watermark from the signal variation sensing unit 400 and the watermark-embedded digital contents, and a high pass filter 540 for extracting only the  
 5 signal information generated between watermarks in the calculated correlation information.

In the above constitution, the watermark restoring filter 410 and the ensemble average calculating unit 520 have the same function as used in the signal variation sensing unit 400 as described above, and thus detailed explanation thereof is omitted.

10 First, in case that the embedded watermark is properly changed according to characteristics of the audio signal and then embedded, water-embedded digital contents undergo the process of the watermark restoration. The expression thereof using an equation is as follows:

$$x(n) = \sum_i b_i sm_j(n-i) \quad (14)$$

15 where,  $sm_j(n)$  indicates a signal of the watermarked digital contents and  $b_i$  indicates a restoration filter coefficient. In order to strengthen the watermark signal, the ensemble average calculating unit 520 removes the periodicity of the digital contents and simultaneously adds the embedded watermark repeatedly. By such calculation, the watermark signal is considered as a main signal, and the digital contents  
 20 signal is considered as a noise. At this time, if the digital contents signal  $s_j(n)$  is random and  $s_j \perp s_i$ , if  $i \neq j$ , the watermark  $x^D(n)$  can be extracted by the ensemble average.

$$x(n) = w(n) + s_j(n) \quad (15)$$

$$x^D(n) = \frac{1}{k} \sum_{j=1}^k [w + s_j] \quad (16)$$

Therefore,  $x^D(n) \approx w(n)$  is established.

The correlation information calculating unit 530 is the same as the correlation information calculating unit 420 in the above Fig. 6. In this case, the correlation information with watermarks is concealed by the periodic characteristics of the digital audio, thus is scarcely discriminated. The present invention provides a method for extracting the correlation information of watermark that is concealed by the periodic characteristics of the digital audio. By using the fact that correlation information has a considerable characteristics of high frequency, while the periodic characteristics of the digital audio has low frequency, only the watermark information can be extracted by a high pass filtering of the correlation information as a hanning window. Such extraction process is presented by the following equation:

$$c(n) = \sum_{k=1}^N x^D(n) w(n+k) \quad (19)$$

$$c^D(n) = \sum_i^N h w(n) c(n+i) \quad (20)$$

The calculated correlation information  $c(n)$  is filtered with the high pass filter  $hw(n)$ , and thus the periodic characteristics of the digital audio contents are excluded and the watermark correlation information is extracted. At the time the above extracted correlation information  $c^D(n)$  is identical to the watermark information, the amount of correlation has quite a higher value than when the correlation amount is not identical.

Fig. 8 shows the constitution of the watermark information authentication unit 600 which extract the information and whether or not the watermark is embedded through  $c^D(n)$  obtained from the watermark detecting unit 500 in Fig. 4. The watermark

information authentication unit 600 in Fig. 8 comprises a peak searching unit 610 for searching a peak in the extracted watermark and a watermark identifying unit 620 for determining whether watermark is embedded or not according to the searched peak.

The peak searching unit 610 of the watermark information authentication unit 5 600 having the constitution described above extracts a particular solution which appears at the time the watermark-embedded digital contents are consistent with watermark. That is, said peak searching unit 610 extracts whether there is a peak having value bigger than a predetermined value and a critical value. The point when a particular solution appears can be predicted since it occurs at a constant interval defined in the 10 embedding process.

In the watermark identifying unit 620, if the extracted particular solution has a value much higher than other value of the correlation information, the digital contents contain the watermark, and if said solution does not have such higher value, the digital contents don't contain watermark. Further, if '10010011' is embedded as watermark 15 information during the embedding process, the positive peak appears when a bit information is '1', and the negative peak appears when a bit information is '0'. Thus, the watermark information of '10010011' can be easily extracted. The capability of extraction with regard to the digital audio contents will be explained by means of the following Embodiments.

20

## COMPARISON OF ABILITY OF DETECTING WATERMARK IN THE DIGITAL AUDIO DATA WITH THE STRONG PERIODICITY

Since the digital audio has periodicity higher than picture signal such as image and video, it is difficult to discriminate whether the watermark exists or not just by the

watermark detecting device already known. It is because the correlation information in watermark is concealed by a strong periodicity which the digital audio signal itself has when the correlation information between the watermark-embedded digital audio and watermark is obtained.

5 Fig. 10A is a result obtained from the existing calculation method of the correlation information, and Fig. 10B is a result obtained from using the watermark perceiving method suggested in the present invention. In case of using the existing calculation method of the correlation coefficient in Fig. 10A, it is impossible to determine whether watermark is embedded due to the periodic characteristics of the  
10 audio, whereas in case of using the watermark detecting method suggested in the present invention, as shown in Fig. 10B, it is possible to detect the watermark with ease since a high value of the correlation coefficient appears at the frequency of the watermark.

## 15 EVALUATION OF ROBUSTNESS OF WATERMARK AGAINST LOSSY COMPRESSION

The digital audio needs the largest amounts of the contents. In particular, it uses many compression algorithm such as ISO/IEC 13818-7 (AA), ISO/IEC 14496-3 (MPEG-4 AAC), ISO/IEC 11172-3 (MP3), Window Media Audio, Twin-VQ, etc. for a  
20 real time transmission through Internet. If watermark is not strong against such lossy compression, it is not possible to utilize watermark for a copyright claim or protection. Thus, watermark should be able to be detected after the lossy compression. In case of using the method for embedding/detecting watermark suggested in the present invention, the watermark is very strong against the lossy compression algorithms known up to now,



and the existence of watermark after the lossy compression apparently appears through the correlation coefficient. Fig. 11 is a result obtained from detecting watermark after the lossy compression.

## 5     **EXAMPLE OF A METHOD OF A EMBEDDING A BIT INFORMATION AND RESTORING INFORMATION**

The decision only as to whether or not watermark exists makes a copyright protection and the chase of illegal copy difficult. Accordingly, a device for embedding/extracting watermark should include a predetermined amount of information  
10    in the digital contents through watermark.

An method for embedding information will be explained with an example of embedding and extracting whether or not a 4bit-information exists. If a '1010' information is embedded during the watermark embedding process as described above when a bit information is '1', watermark is embedded to have the correlation of the  
15    positive, and when a bit information is '0', watermark is embedded to have the correlation of the negative. Such repeated embedding of watermark enables longer information to be easily obtained. Further, in case of embedding a long information, it is necessary to understand the order of information. Since information is not always extracted in the same order, watermark produced by using a different value of key is  
20    embedded into the  $k$  th signal, thereby the order can be easily prevented from being disarranged.

For example, when the watermark produced as a key value of '1234' is embedded, and the first bit information '1' produced as a key value of '1235' is embedded, only this signal is discriminated as watermark differently produced.

Otherwise, it enables the starting point of information to be found by adjusting the size of the strength of watermark. Fig. 12 shows a result of extraction when embedding a 4-bit information. It can be understood that the first peak from the right in the above-mentioned figure is a starting point.

5

### INDUSTRIAL APPLICABILITY

A psycho-acoustic model is formed in order not to affect the audio characteristic of the digital audio when embedding a digital watermark in the digital audio contents according to the present invention as described above, and an audio absolute threshold of hearing is used in filtering. Thus, it prevents the signal in which a watermark is not embedded from being discriminated by audibly changing the quality of the contents. Further, since the psycho-acoustic model is made by using person's splitting characteristics with respect to audio, the embedded information enables a deterioration of quality of the audio to be minimized when a watermark is embedded in the digital audio contents.

Further, the psycho-acoustic model has a characteristic strong in processing a digital signal while the digital contents and the embedded watermark signal maintain conservation of the digital contents and quality of the contents. Such characteristics can help detect the watermark by a component of the other parts, even if a part of the frequency component is removed since the volume of signal is adjusted according to the characteristics of the audio frequency when masking the pseudo-random number sequence by using the psycho-acoustic model.

Further, when extracting the embedded watermark according to the present invention, watermark can be successfully extracted in a case where a part of the

watermark disappears. The watermark included in the digital audio contents can be easily extracted by strengthening the watermark signal during the process.

As described above, the present invention provides a method for extracting watermark embedded from the signal with a strong periodicity such as the digital audio  
5 in a high-speed manner and in effectiveness. It is possible to extract the embedded information by such extraction method as long as deterioration in the quality of the audio contents maintains the commercial value of audio even after the attack of an audio signal processing (analog to digital transformation, transformation in a sampling ratio, transformation in a linear speed, lossy compression, echo hiding).

10 The present invention is particularly shown and described referring to the above Embodiments, however, which are used for examples. An ordinarily skilled person in the art to which the present invention pertains can make various revisions without deviating beyond the mind and scope of the invention as defined in the claims attached hereto.

**What is claimed is:**

1. A method for embedding a digital watermark in the digital audio contents, comprising the steps of:

5       generating a digital watermark by filtering a predetermined length of pseudo random sequence with audio absolute threshold of hearing in order not to affect the audio characteristics of said digital audio contents; and

          embedding said digital watermark in said digital audio contents.

10    2. The method according to claim 1, wherein said audio absolute threshold of hearing is the minimum level of the audio perceivable by hearing meaning the limit of noise perceivable by hearing when said digital audio contents are reproduced.

3. The method according to claim 1, wherein the step of generating said digital  
15    watermark includes the steps of:

          forming a psycho-acoustic model having the audio absolute threshold of hearing according to the characteristics of the frequency generated from said digital audio contents;

          filtering said pseudo random number sequence in accordance with said psycho-  
20    acoustic model; and

          filtering said filtered pseudo random number sequence by using composed filtering coefficient which indicates the general characteristics of said digital audio contents.

4. The method according to claim 3, wherein the size of filtering said pseudo random

number sequence according to said psycho-acoustic model is adjusted according to the general characteristics of frequency of said digital audio contents.

- 5 5. The method according to claim 1, wherein the step of embedding said digital watermark in said digital audio contents includes the steps of:

measuring the amount of energy of said digital audio contents;

adjusting the strength of said digital watermark according to said measured amount of energy; and

- 10 adding up said digital watermark with adjusted strength with said digital audio contents.

6. The method according to claim 5, wherein said measurement of the amount of energy is accomplished by using the volume of said digital audio contents.

- 15 7. The method according to claim 5, wherein said measurement of the amount of energy is accomplished by using the characteristics of frequency distribution in said digital audio contents.

- 20 8. A method for extracting a watermark from the digital audio contents, the method including the steps of:

generating a watermark from pseudo random number sequence by forming a psycho-acoustic model having an audio absolute threshold of hearing according to the characteristics of frequency generated from said digital audio contents; and

extracting the embedding of watermark by measuring the correlation between

said watermark and said digital audio contents.

9. The method according to claim 8, further including the step of adjusting the length of said watermark by detecting whether or not the signal is processed with regard to the  
5 corresponding digital audio contents from the digital audio contents which are adjacent to each other after said step of generating watermark.

10. The method according to claim 9, wherein the step of generating said watermark includes the steps of:

10 forming a psycho-acoustic model having audio absolute threshold of hearing according to the characteristics of frequency generated from said digital audio contents; and

generating said watermark by filtering said pseudo random number sequence in accordance with said psycho-acoustic model.

15

11. The method according to claim 10, wherein the step of adjusting the length of said watermark includes the steps of:

restoring back to its original form regarding the signal processed when embedding watermark is performed regarding said adjacent digital audio contents;

20 extracting the information of correlation among said restored adjacent digital audio contents; and

resampling said watermark according to said information of correlation.

12. The method according to claim 11, further including the step of strengthening said

information of correlation after said step of extracting .

13. The method according to claim 11, wherein said step of restoring is conducted by inverse-auto regression filter.

5

14. The method according to claim 12, wherein said step of strengthening the information of correlation performs the calculation of the ensemble average.

15. The method according to claim 11, wherein said step of extracting said  
10 embedding of watermark includes the steps of:

restoring back to its original form regarding the signal processed when embedding watermark is performed in said adjacent digital audio contents;

strengthening the component of said digital watermark among said restored digital audio contents;

15 calculating the information of correlation between two signals of said watermark with adjusted length and said digital audio contents with strengthened digital watermark; and

extracting the information of watermark correlation by filtering said calculated and extracted correlation information to remove the periodic property of said digital  
20 audio contents.

16. The method according to claim 15, further including the step of determining whether watermark is embedded from said correlation information of the extracted watermark after said step of extracting watermark.

17. A device for embedding watermark in the digital audio contents, the device including :

- a means for generating the digital watermark by filtering a predetermined length of pseudo random sequence with audio absolute threshold of hearing in order not to affect the audio characteristics of said digital audio contents; and
- a means for embedding said digital watermark in said digital audio contents.

18. The device according to claim 17, wherein the means for generating said digital watermark includes:

- a means for forming a psycho-acoustic model having the audio absolute threshold of hearing from analyzing the characteristics of frequency generated from said digital audio contents;
- a means for filtering said pseudo random number sequence in accordance with said psycho-acoustic model; and
- a means for filtering said filtered pseudo random number sequence by using composed filtering coefficient which indicates the general characteristics of said digital audio contents.

19. The device according to claim 18, wherein the strength of filtering said pseudo random number sequence in accordance with said psycho-acoustic model is adjusted according to the general characteristics of frequency of said digital audio contents.

20. The device according to claim 17, wherein the means for embedding said digital watermark



includes:

a means for measuring the amount of energy of said digital audio contents;

a means for adjusting the strength of said digital watermark according to said measured amount of energy; and

5 a means for adding up said digital watermark with adjusted strength with said digital audio contents.

21. The device according to claim 20, wherein said means for measuring the amount of energy measures the volume of said digital audio contents.

10

22. The device according to claim 20, wherein said means for measuring the amount of energy measures the characteristics of frequency distribution of said digital audio contents.

15 23. A device for extracting watermark from the digital audio contents, the device including:

a means for generating a watermark by forming a psycho-acoustic model having an audio absolute threshold of hearing from analyzing the characteristics of frequency generated from said digital audio contents; and

20 a means for extracting the embedding of watermark by measuring the correlation between said watermark and said digital audio contents.

24. The device according to claim 23, further including the means of adjusting the length of said watermark by detecting whether or not the signal is processed with regard

to the corresponding digital audio contents from the digital audio contents which are adjacent to each other, and said watermark with adjusted length is input to said extracting means.

- 5 25. The device according to claim 24, wherein said means for adjusting the length of the watermark includes:

a means for restoring back to its original form regarding the signal processed when embedding watermark is performed regarding said adjacent digital audio contents;

- 10 a means for extracting the information of correlation among said restored adjacent digital audio contents; and

a means for resampling said watermark according to said information of correlation.

- 15 26. The device according to claim 25, further including the means for strengthening said information of correlation obtained by said extracting means.

27. The device according to claim 24, wherein said restoring means is an inverse auto-regression filtering.

- 20 28. The device according to claim 26, wherein said means for strengthening the information of correlation performs the calculation of the ensemble average.

29. The device according to claim 25, wherein said means for extracting the embedding of watermark includes:

a means for restoring back to its original form regarding the signal processed when embedding watermark is performed regarding said adjacent digital audio contents;

a means for strengthening the component of said digital watermark among said restored digital audio contents;

5 a means for calculating the information of correlation between two signals of said watermark with adjusted length, and said digital audio contents with strengthened digital watermark; and

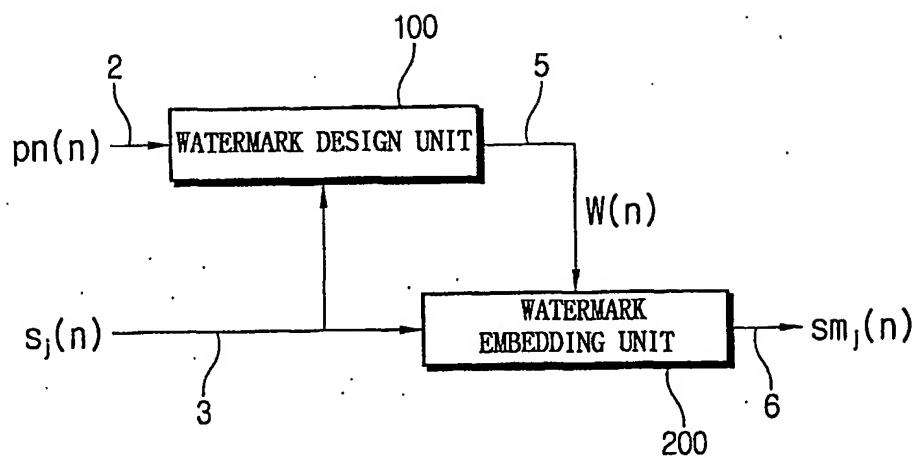
a means for extracting the information of watermark correlation by filtering said calculated and extracted correlation information to remove the periodic property of  
10 said digital audio contents.

30. The device according to claim 29, further including the means for determining whether or not watermark is embedded from said correlation information of said extracted watermark.

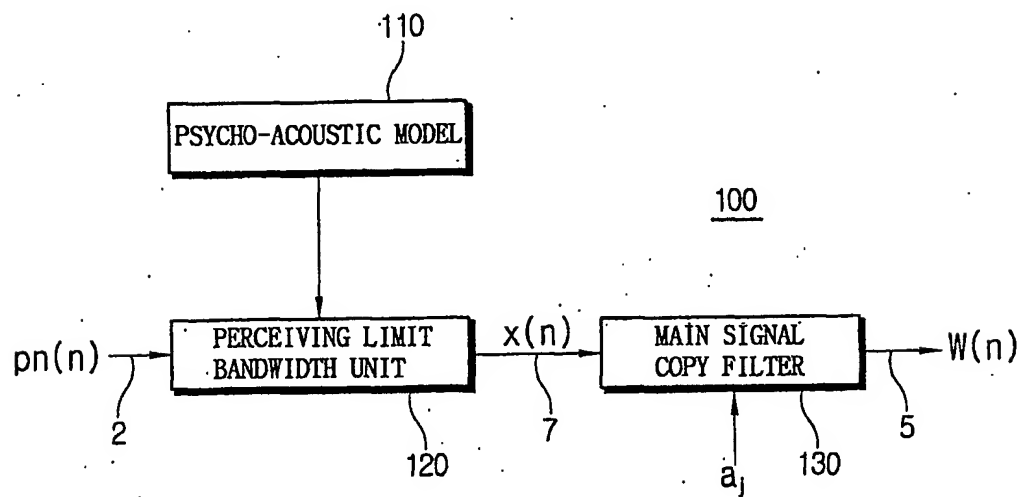
15

1/7

FIG. 1

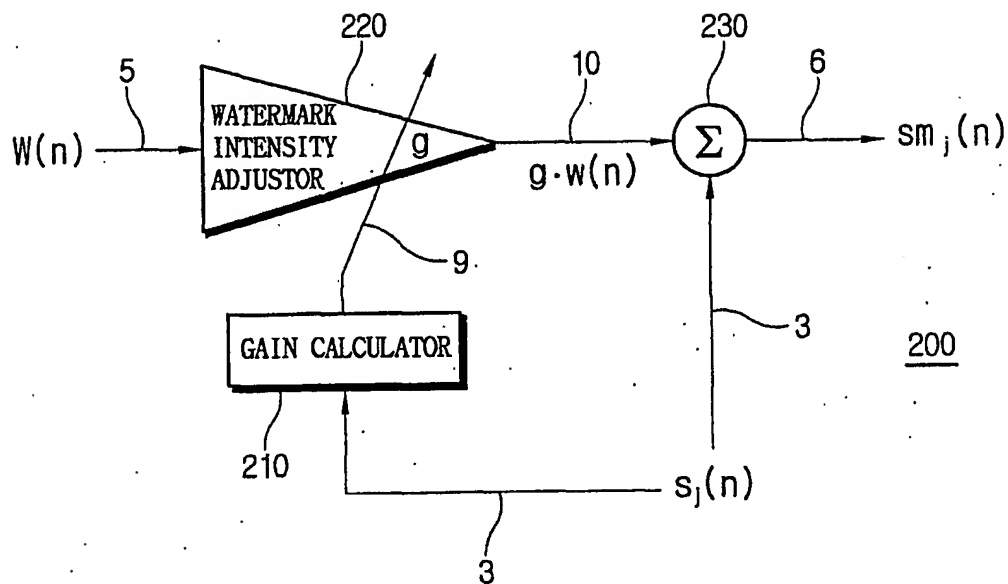


5 FIG. 2

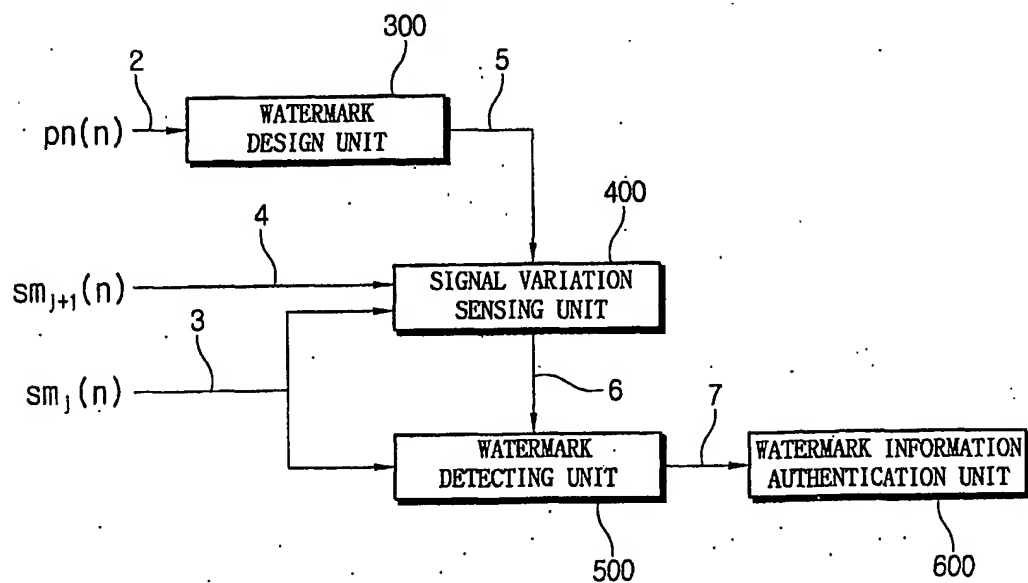


2/7

FIG. 3

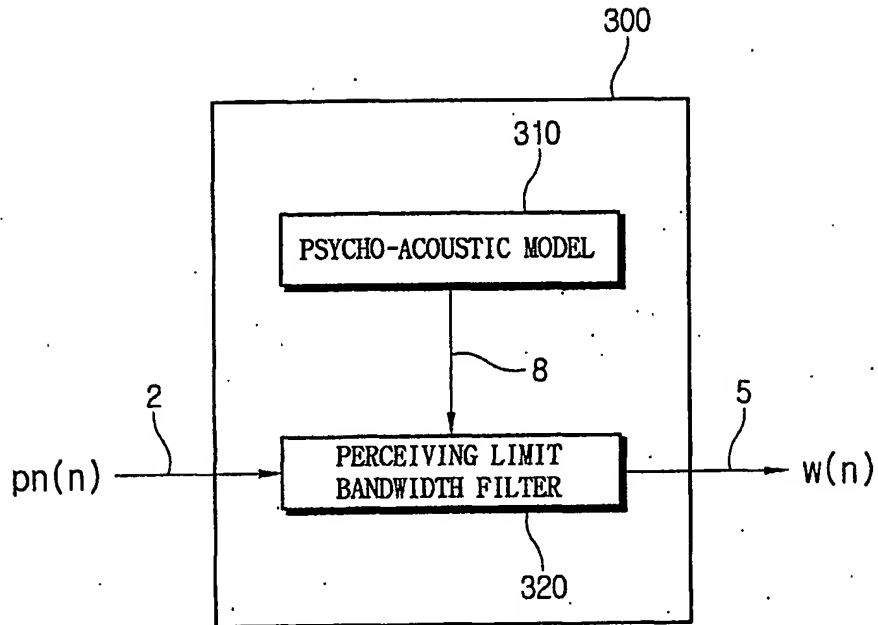


5 FIG. 4

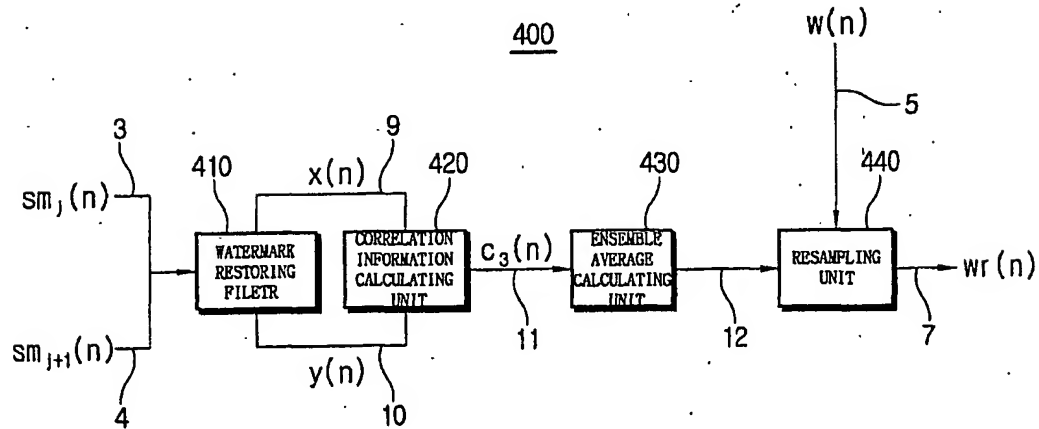


3/7

FIG. 5

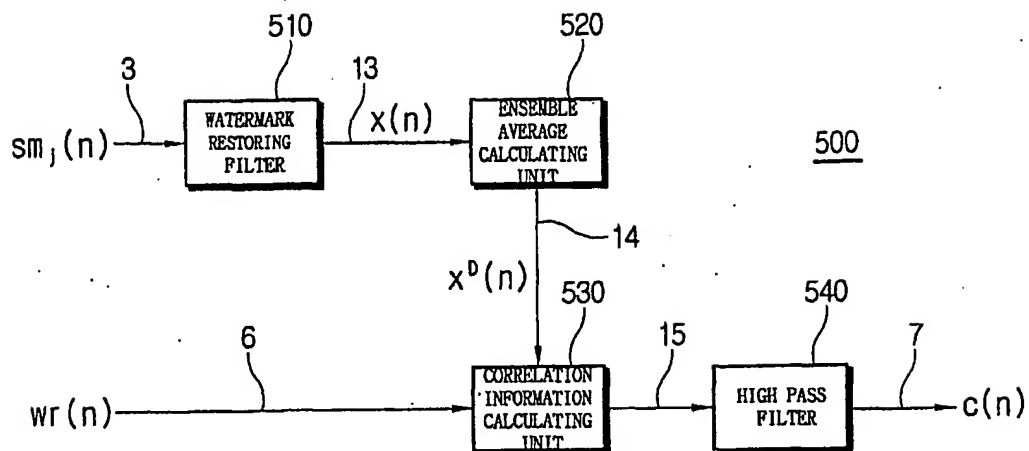


5 FIG. 6

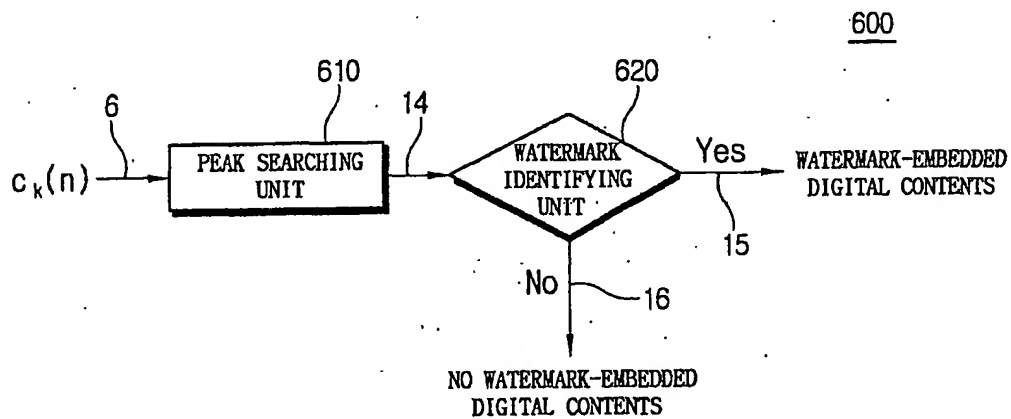


4/7

FIG. 7



5 FIG. 8



5/7

FIG. 9

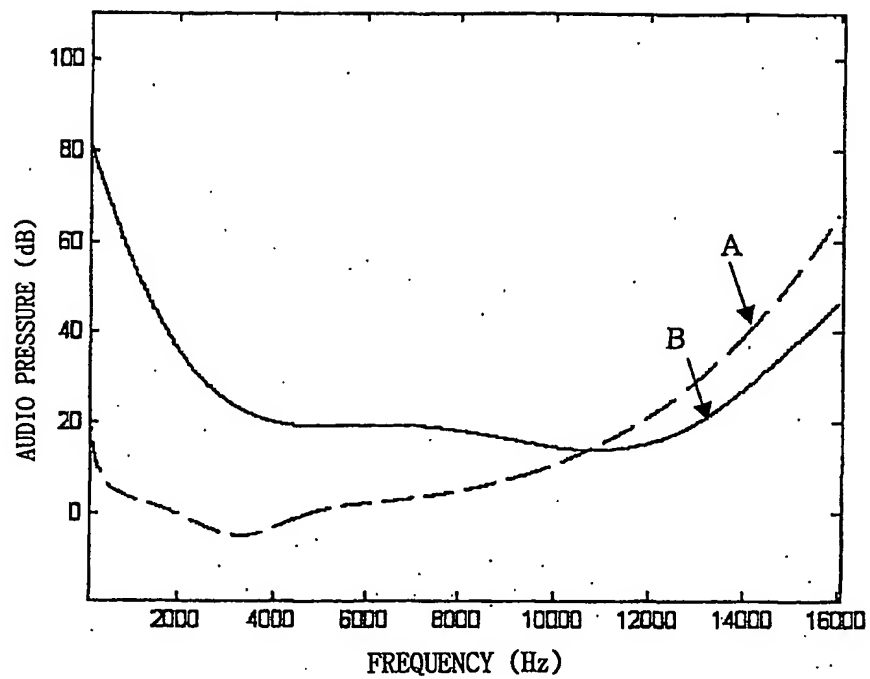
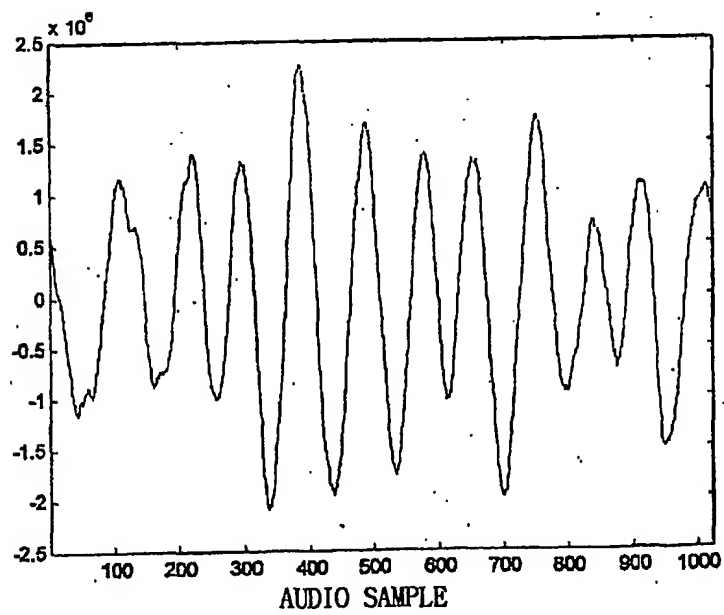


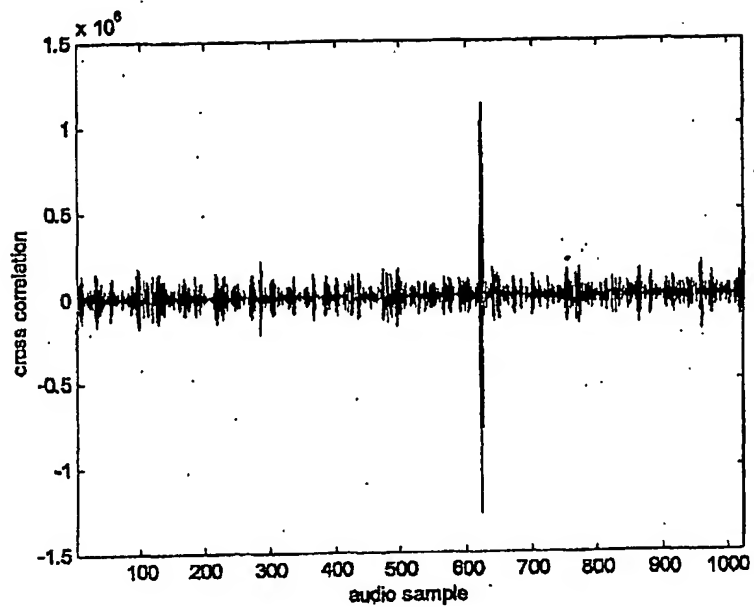
FIG. 10A



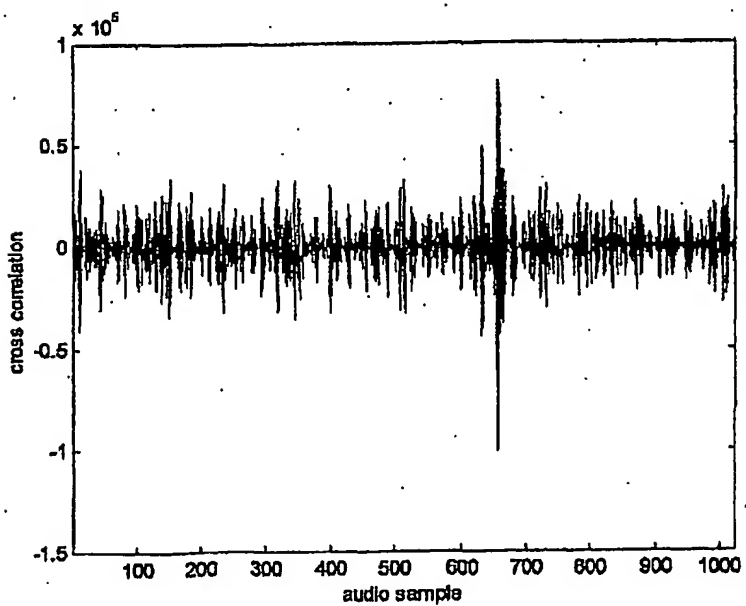


6/7

FIG. 10B

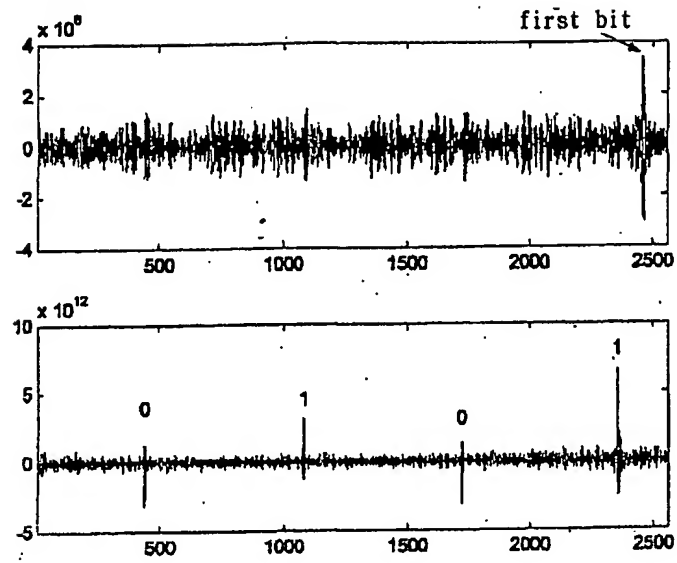


5 FIG. 11



7/7

FIG. 12



## INTERNATIONAL SEARCH REPORT

 International application No.  
PCT/KR01/00975
**A. CLASSIFICATION OF SUBJECT MATTER****IPC7 G11B 20/10**

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

IPC7 G11B20/10 G06F17/30 H04L9/32 G09C5/00

 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
Korean Patents and Applications for Inventions since 1975

 Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
WPI, PAJ "WATERMARK""EMBEDDING""DIGITAL""STREAM"
**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5,737,495 (INTEL CO) 7 APRIL 1998 see the whole document	1, 8, 17, 23
A	US 5,765,152 (TRUSTEES OF DARTMOUTH COLLEGE) 9 JUNE 1998 see the whole document	1, 8, 17, 23
A	US 5,910,987 (InterTrust Co) 8 JUNE 1999 see the whole document	1, 8, 17, 23

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

31 OCTOBER 2001 (31.10.2001)

Date of mailing of the international search report

31 OCTOBER 2001 (31.10.2001)

Name and mailing address of the ISA/KR

 Korean Intellectual Property Office  
Government Complex-Daejeon, Dunsan-dong, Seo-gu, Daejeon  
Metropolitan City 302-701, Republic of Korea

Facsimile No. 82-42-472-7140

Authorized officer

SEO, Hawthorne

Telephone No. 82-42-481-5693

